

Focusing in Arthurs-Kelly-type Joint Measurements with Correlated Probes

Thomas J Bullock* and Paul Busch†

Department of Mathematics, University of York, York, UK

(Dated: September 17, 2014)

Joint approximate measurement schemes of position and momentum provide us with a means of inferring pieces of complementary information if we allow for the irreducible noise required by quantum theory. One such scheme is given by the Arthurs-Kelly model, where information about a system is extracted via indirect probe measurements, assuming separable uncorrelated probes. Here, following Di Lorenzo (PRL **110**, 120403 (2013)), we extend this model to both entangled and classically correlated probes, achieving full generality. We show that correlated probes can produce more precise joint measurement outcomes than the same probes can achieve if applied alone to realize a position or momentum measurement. This phenomenon of *focusing* may be useful where one tries to optimize measurements with limited physical resources. Contrary to Di Lorenzo's claim, we find that there are no violations of Heisenberg's error-disturbance relation in these generalized Arthurs-Kelly models. This is simply due to the fact that, as we show, the measured observable of the system under consideration is covariant under phase space translations and as such is known to obey a tight joint measurement error relation.

Introduction. The incompatibility of the position and momentum observables is a well-known feature of quantum mechanics and is succinctly expressed by the preparation uncertainty relation [1, 2]:

$$\text{Var}(Q, \psi)\text{Var}(P, \psi) \geq \frac{\hbar^2}{4}, \quad (1)$$

which states that for any state ψ that we prepare a system in, the product of the variances in the statistics of the position Q and momentum P is bounded below by Planck's constant. However, while (1) is well understood as a preparation uncertainty relation, there is controversy over how this incompatibility may be expressed when we consider measurements of *both* observables on the same system. Heisenberg [3] formulated a trade-off relation for the error ΔQ of a position measurement and the resulting disturbance ΔP of momentum:

$$\Delta Q \Delta P \geq \frac{\hbar}{2}, \quad (2)$$

which he obtained on the basis of heuristic arguments. In recent years this form of tight bound has been called into question by some [1, 5, 6] and corroborated by others [2, 8, 9]. This controversy is the result of a lack of universally agreed upon operational definitions of error and disturbance.

Here we analyze the work of Di Lorenzo [6], whose claim of a violation of (2) results from the consideration of a particular measurement model and a specific choice of measures of error and disturbance. His scheme, a generalization of the celebrated model of Arthurs and Kelly [12], couples a system to two probes that are then measured to provide approximate information about both the position and momentum of the system. The suggested extension consists of allowing for initial correlations between the two probes. We show that the purported violation of (2) does not occur in the most general extension of the Arthurs-Kelly model.

In what follows, we consider the generalized Arthurs-Kelly model and allow for correlation between the probes. We derive the effective joint observable measured on the system, which is represented operationally as a positive operator valued measure (POVM) on phase space. It is found that, for arbitrary preparations of the probes, this observable is covariant under phase space translations; i.e., such translations do not change the observable but instead shift its associated probability distributions. This generalizes the case of uncorrelated probes, where the covariance property has previously been shown [10].

Covariant phase space observables are known to always satisfy an error-error relation of the form (2) [2]. Given that a sequential measurement is a form of joint measurement, with the disturbance corresponding to the approximation error in the second observable (see, e.g., [1, 2, 11]), it follows that error-disturbance relations are special cases of error-error relations. That the error and disturbance measures given in [6] lead to a violation of (2) suggests something is wrong with these measures.

In fact, we show that Di Lorenzo's "disturbance" is actually a measure of the relative imprecision of two approximations of the ideal position or momentum observable. The first approximation is the marginal observable derived from the Arthurs-Kelly model, while the second is the one measured by a single probe. As Di Lorenzo found, this relative imprecision can indeed become negative. This observation leads to an interesting phenomenon that we refer to as *focusing*: the marginal observables in a joint measurement can be more precise than the observables measured by the individual probes. Focusing may prove useful for improving the performance of measurements with limited physical resources.

We give two cases where both the approximate position and momentum observables are focused by performing a joint measurement. In one instance the probes are prepared in a pure entangled state, while in the second

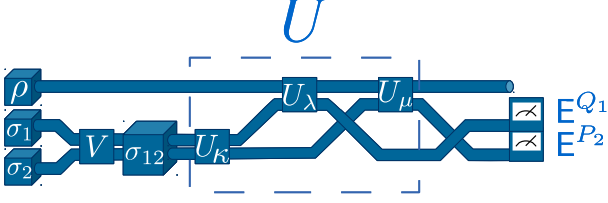


FIG. 1. An extension of the Arthurs-Kelly model to allow for correlated states. Two probes, which are coupled together by a unitary V , are then coupled to the system described by state ρ by a unitary U . After this coupling, ideal measurements of the position of the first probe and momentum of the second are performed, from which we infer information about the position and momentum of our considered system.

they are prepared in a separable mixed state. This latter case shows that entanglement does not help over classical correlations.

In what follows, we will consider three particles with one continuous degree of freedom. Each particle is described by either a pure state belonging to the Hilbert space $L^2(\mathbb{R})$ of square-integrable complex functions over the real line \mathbb{R} or by a density operator. Further to this, we will set $\hbar = 1$ for simplicity.

The model. We consider an extension of the Arthurs-Kelly model [12], as shown in Fig. 1. The model couples a system, described by a state ρ , to two probes, labeled 1 and 2, via an impulsive (short-timed) unitary interaction $U = \exp(-iH)$, where H is the interaction Hamiltonian

$$H = \lambda Q P_1 - \mu P Q_2 + \frac{\lambda\mu}{2} \kappa P_1 Q_2. \quad (3)$$

The numbered operators P_1 and Q_2 refer to the momentum and position operators on probes 1 and 2, respectively, while the unnumbered operators Q and P are the position and momentum operators, respectively, on the considered system. The positive coupling constants λ and μ determine the strength of the coupling between the two probes and the system, while κ determines the coupling strength between the two probes. Using Baker-Campbell-Hausdorff decompositions of U :

$$\begin{aligned} U &= \exp(i\mu P Q_2) \exp(-i\lambda Q P_1) \exp[-i\frac{\lambda\mu}{2}(\kappa - 1)P_1 Q_2] \\ &= \exp(-i\lambda Q P_1) \exp(i\mu P Q_2) \exp[-i\frac{\lambda\mu}{2}(\kappa + 1)P_1 Q_2], \end{aligned} \quad (4)$$

we see that the joint measurement can alternatively be implemented as a sequence of interactions and measurements. If $|\kappa| = 1$, this can be considered as a strictly sequential measurement of position and momentum (with the ordering depending on the sign of κ). These coupling constants are assumed large enough that we may ignore the free evolution of the system and the probes. After this coupling, the two probes are ideally measured, i.e., we perform projection-valued measures on the probes. The first probe has its position measured by E^{Q_1} , and

the second probe its momentum by E^{P_2} . From the statistics of these measurements, we infer information about the position and momentum of the considered system. In most works, the probes were assumed to be in pure, uncorrelated states; i.e., they are described by a product state $\varphi_1 \otimes \varphi_2$, say, with φ_i the state of the i th probe. Assuming such a situation for the probes, the model produces an effective joint observable G on the considered system that is covariant under phase space translations [10]. By covariance we mean that the application of a phase space translation $W_{qp} = \exp[-i(qP - pQ)]$ to the POVM element $G(Z)$, with $Z \subseteq \mathbb{R}^2$ an interval in phase space, will result in another element of the same POVM, but with a shifted input value $Z + (q, p)$:

$$W_{qp} G(Z) W_{qp}^* = G(Z + (q, p)). \quad (5)$$

Instead of the pure product state $\varphi_1 \otimes \varphi_2$ describing the probes, we consider a generally correlated state σ_{12} , which may be seen as the result of a unitary coupling V ; i.e., $\sigma_{12} = V(\sigma_1 \otimes \sigma_2)V^*$, as in Fig. 1 [13]. In the case σ_{12} is a pure entangled state, $\sigma_{12} = P_{\varphi_{12}}$ (the projector onto the normalized vector φ_{12}), we find coupling the probes to our system initially in a pure state $\rho = P_\psi$ produces the state of the combined system in the position representation:

$$\begin{aligned} \Psi(q, q_1, q_2) &= U(\psi \otimes \varphi_{12})(q, q_1, q_2) \\ &= \psi(q + \mu q_2) \varphi_{12}(q_1 - \lambda q - \frac{\lambda\mu}{2}(\kappa + 1)q_2, q_2), \end{aligned} \quad (6)$$

where q , q_1 and q_2 are the position coordinates for the system and probes 1 and 2, respectively. Similarly, if σ_{12} is mixed, we consider the mixed state $U(P_\psi \otimes \sigma_{12})U^*$. From this, the probability of finding the pointer readings belonging to the intervals $X, Y \subseteq \mathbb{R}$, respectively, is given by $\langle \Psi | I \otimes E^{Q_1}(X) \otimes E^{P_2}(Y) | \Psi \rangle$. With appropriate scaling of the intervals, this probability can be interpreted as the likelihood of finding the system, described by state ψ , belonging to the phase space cell $X \times Y \subseteq \mathbb{R}^2$:

$$\langle \Psi | I \otimes E^{Q_1}(\lambda X) \otimes E^{P_2}(\mu Y) | \Psi \rangle \equiv \langle \psi | G^{(\lambda, \mu)}(X \times Y) | \psi \rangle. \quad (7)$$

(Here $\lambda X = \{\lambda x | x \in X\}$, and similarly for μY .) The positive operators $G^{(\lambda, \mu)}(X \times Y)$ are mathematically defined by (7) and form a phase space observable on the system. We can now state our main result [15].

Theorem 1. *The observable $G^{(\lambda, \mu)}$ given by an Arthurs-Kelly model with arbitrary probe state is a covariant phase space observable.*

This extends what has been considered in the past to all possible probe preparations for the Arthurs-Kelly model and shows that the results found for probes prepared in pure product states can be readily generalized.

Covariant phase space observables have been studied thoroughly elsewhere (see, e.g., [18–21]), and it is well

known that any such observable G may be expressed in the form

$$G(Z) = G_\tau(Z) = \frac{1}{2\pi} \int_Z dq dp W_{qp} \tau W_{qp}^*, \quad (8)$$

where $Z \subseteq \mathbb{R}^2$ and τ is a unique positive operator with unit trace; i.e., mathematically, τ is a density operator. The marginal observables of G are approximations of the ideal position and momentum operators. In particular, the marginals of $G^{(\lambda,\mu)}$, denoted by $E^{(\lambda,\mu)}$ and $F^{(\lambda,\mu)}$, are

$$E^{(\lambda,\mu)}(X) = G^{(\lambda,\mu)}(X \times \mathbb{R}) = (\chi_X * e^{(\lambda,\mu)})(Q), \quad (9a)$$

$$F^{(\lambda,\mu)}(Y) = G^{(\lambda,\mu)}(\mathbb{R} \times Y) = (\chi_Y * f^{(\lambda,\mu)})(P), \quad (9b)$$

where χ_A denotes the indicator function onto the subset $A \subseteq \mathbb{R}$, $*$ denotes convolution, and both $e^{(\lambda,\mu)}$ and $f^{(\lambda,\mu)}$ are probability distributions. These probability distributions characterize the noise in the measurement statistics of $E^{(\lambda,\mu)}$ and $F^{(\lambda,\mu)}$ and depend on the state σ_{12} of the probes and the coupling constants. Owing to the structure of G as given in (8), these distributions are identical to the probability distributions of position and momentum in the state represented by τ_- , the space inversion of τ . Hence it is evident that their variances obey the standard uncertainty relation

$$\text{Var}(e^{(\lambda,\mu)}) \text{Var}(f^{(\lambda,\mu)}) \geq \frac{1}{4}. \quad (10)$$

The second moment of (say) the distribution $e^{(\lambda,\mu)}$ is expressed in terms of the first moment (the mean) and the variance via

$$e^{(\lambda,\mu)}[2] = e^{(\lambda,\mu)}[1]^2 + \text{Var}(e^{(\lambda,\mu)}). \quad (11)$$

This can be interpreted as representing both systematic and random error contributions inherent in $E^{(\lambda,\mu)}$ as an approximation of the ideal position measurement E^Q . This intuition is strengthened by the fact that one can find physically relevant measures of the error, Δ , of approximating an ideal observable by another, such that $\Delta(E^{(\lambda,\mu)}, E^Q)^2 = e^{(\lambda,\mu)}[2]$ [15]. The inequality

$$e^{(\lambda,\mu)}[2] f^{(\lambda,\mu)}[2] \geq 1/4, \quad (12)$$

obtained as a direct consequence of (10), is therefore an instance of the joint measurement uncertainty relation (2). This result holds for all covariant phase space observables, including the ones that arise from our extension of the Arthurs-Kelly model.

As we shall show, Di Lorenzo's claim of a violation of (2) is a result of an inadequate definition of disturbance.

Di Lorenzo's disturbance. We first consider the model of the previous section and set one of the coupling constants, either λ or μ , to zero. If we put $\mu = 0$, the coupling unitary U reduces to $U_\lambda = \exp(-i\lambda Q P_1)$, and

we measure with just the first probe. Similarly, if we set $\lambda = 0$, then U reduces to $U_\mu = \exp(i\mu P Q_2)$, and we measure with the second probe. The first instance results in the effective observable $E^{(\lambda,0)}$, and the second $F^{(0,\mu)}$. These observables have a form similar to those in (9a) and (9b):

$$\begin{aligned} E^{(\lambda,0)}(X) &= (\chi_X * e^{(\lambda,0)})(Q), \\ F^{(0,\mu)}(Y) &= (\chi_Y * f^{(0,\mu)})(P). \end{aligned} \quad (13)$$

By using Eqs. (9a), (9b) and (13), the disturbance given by Di Lorenzo Δ_{DL} is the difference in the variances of the measurement statistics for the marginal observable and its individual measurement counterpart with regard to some system state ψ , i.e.,

$$\begin{aligned} \Delta_{DL}(Q) &= \text{Var}(E^{(\lambda,\mu)}, \psi) - \text{Var}(E^{(\lambda,0)}, \psi) \\ &= \text{Var}(e^{(\lambda,\mu)}) - \text{Var}(e^{(\lambda,0)}), \end{aligned} \quad (14a)$$

$$\begin{aligned} \Delta_{DL}(P) &= \text{Var}(F^{(\lambda,\mu)}, \psi) - \text{Var}(F^{(0,\mu)}, \psi) \\ &= \text{Var}(f^{(\lambda,\mu)}) - \text{Var}(f^{(0,\mu)}). \end{aligned} \quad (14b)$$

Note that the state-dependent parts vanish due to the additive nature of the variance of convolutions. If we use a general correlated state σ_{12} for our probes, then these errors take the form

$$\begin{aligned} \Delta_{DL}(Q) &= \frac{(1-\kappa)^2}{4} \mu^2 \text{Var}(Q_2, \sigma_{12}) \\ &\quad - \frac{(1-\kappa)}{\lambda} \mu \text{Cov}(Q_1, Q_2, \sigma_{12}), \end{aligned} \quad (15)$$

$$\begin{aligned} \Delta_{DL}(P) &= \frac{(1+\kappa)^2}{4} \lambda^2 \text{Var}(P_1, \sigma_{12}) \\ &\quad - \frac{(1+\kappa)}{\mu} \lambda \text{Cov}(P_1, P_2, \sigma_{12}), \end{aligned} \quad (16)$$

where $\text{Cov}(A, B, \rho)$ denotes the covariance between the two observables with regards to the state ρ : $\text{Cov}(A, B, \rho) = \text{tr}[\frac{1}{2}(AB + BA)\rho] - \text{tr}[A\rho] \text{tr}[B\rho]$.

If we begin with the states of our probes being uncorrelated, i.e., $\sigma_{12} = \sigma_1 \otimes \sigma_2$, then the covariance terms in $\Delta_{DL}(Q)$ and $\Delta_{DL}(P)$ vanish, and the measures are strictly non-negative. However, the inclusion of correlated probe states in the model means that the covariance terms can be nonzero and, indeed, can in some instances be large enough that Δ_{DL} is negative. Interpreting this negativity as indicating the absence of disturbance would be implausible: *any* nonzero value of Δ_{DL} indicates an influence of the other measurement. The occurrence of negative values is not surprising, however, since Di Lorenzo begins by calibrating with a poor reference measurement; had he chosen the ideal reference measurement, then his disturbance value would coincide with the random error contribution in (11), and he would have recovered (10).

By using correlated probe states we introduce the concept of focusing, where $E^{(\lambda,\mu)}$ ($F^{(\lambda,\mu)}$) is a more precise approximation of position (momentum) than $E^{(\lambda,0)}$

$(F^{(0,\mu)})$, despite being the marginal of a joint observable. Note that all that matters for this focusing to be able to occur is the existence of some initial correlation between the probes, and in this sense entanglement is no more significant than being able to prepare the probes in a mixed state.

In the next section we show two examples in which our model can lead to both $\Delta_{\text{DL}}(Q)$ and $\Delta_{\text{DL}}(P)$ being negative simultaneously. In such cases we have a setup in which a joint measurement of position and momentum is more precise than if we separately performed individual measurements as described above. In what follows, we relabel Δ_{DL} by \mathcal{F} , so $\mathcal{F}(Q) := \Delta_{\text{DL}}(Q)$, etc.

Examples of focusing. The first case we will consider is that of probes prepared in a pure entangled state [22]. In particular, we choose the unbiased two-mode Gaussian state, given in the position representation by

$$\varphi_{12}(q_1, q_2) = \left(\frac{4 \det D}{\pi^2} \right)^{1/4} \exp[-(q_1, q_2) D (q_1, q_2)^T], \quad (17)$$

where D is the positive-definite matrix

$$D = \begin{pmatrix} a & b \\ b & d \end{pmatrix} = \begin{pmatrix} \langle P_1^2 \rangle_{\varphi_{12}} & \langle P_1 P_2 \rangle_{\varphi_{12}} \\ \langle P_1 P_2 \rangle_{\varphi_{12}} & \langle P_2^2 \rangle_{\varphi_{12}} \end{pmatrix} \quad (18)$$

$$= 4 \det D \begin{pmatrix} \langle Q_2^2 \rangle_{\varphi_{12}} & -\langle Q_1 Q_2 \rangle_{\varphi_{12}} \\ -\langle Q_1 Q_2 \rangle_{\varphi_{12}} & \langle Q_1^2 \rangle_{\varphi_{12}} \end{pmatrix}.$$

Here we have used the shorthand $\langle A \rangle_\psi = \langle \psi | A | \psi \rangle$. Since φ_{12} is unbiased, i.e., $\langle Q_i \rangle_{\varphi_{12}} = \langle P_i \rangle_{\varphi_{12}} = 0$ for $i = 1, 2$, the variances and covariances reduce: $\text{Var}(Q_i, \varphi_{12}) = \langle Q_i^2 \rangle_{\varphi_{12}}$ and $\text{Cov}(Q_1, Q_2, \varphi_{12}) = \langle Q_1 Q_2 \rangle_{\varphi_{12}}$, etc. With this in mind, Eqs. (15) and (16) can be expressed in terms of the components of D , and so the conditions $\mathcal{F}(Q) < 0$ and $\mathcal{F}(P) < 0$ are equivalent to

$$-\frac{(1-\kappa)}{\lambda} \mu b > \frac{(1-\kappa)^2}{4} \mu^2 a > 0, \quad (19)$$

$$\frac{(1+\kappa)}{\mu} \lambda b > \frac{(1+\kappa)^2}{4} \lambda^2 a > 0. \quad (20)$$

If we set $|\kappa| < 1$, then both $1 + \kappa$ and $1 - \kappa$ are positive, so from Eq. (19) we find that $b < 0$, while Eq. (20) shows that $b > 0$. (In order to arrive at these inequalities we have used the fact that $\lambda, \mu > 0$.) We conclude that, while using the two-mode Gaussian state φ_{12} , we are unable to achieve both $\mathcal{F}(Q) < 0$ and $\mathcal{F}(P) < 0$ if we set $|\kappa| < 1$.

If, however, $|\kappa| > 1$, then it is possible to attain both $\mathcal{F}(Q) < 0$ and $\mathcal{F}(P) < 0$. By considering the cases $\kappa > 1$ and $\kappa < -1$ separately, it is quickly shown that we can achieve focusing on both approximate observables for the state φ_{12} iff $|\kappa| > 1$ and

$$\frac{\lambda \mu}{4} (1 + |\kappa|) < \frac{|b|}{a}. \quad (21)$$

The second example that we consider is a mixed state $\sigma_{12} = \sigma = p\sigma_1 + (1-p)\sigma_2$, where $0 < p < 1$. Both pure states σ_i are product states composed of two single-mode Gaussian states centered on the point (x_i, k_i) in phase space, i.e., $\sigma_i = P_{\varphi_i^{(1)} \otimes \varphi_i^{(2)}}$ where $\langle Q_1 \rangle_{\sigma_i} = \langle Q_2 \rangle_{\sigma_i} = x_i$ and $\langle P_1 \rangle_{\sigma_i} = \langle P_2 \rangle_{\sigma_i} = k_i$. Further to this, we assume that the pure states have a fixed variance S with respect to both position operators and R with respect to both momentum operators, i.e. $\text{Var}(Q_1, \sigma_i) = \text{Var}(Q_2, \sigma_i) = S$ and $\text{Var}(P_1, \sigma_i) = \text{Var}(P_2, \sigma_i) = R$ for both i . The covariances of Q_1 , Q_2 and P_1 , P_2 with respect to the state σ are then

$$\begin{aligned} \text{Cov}(Q_1, Q_2, \sigma) &= (p - p^2)(x_1 - x_2)^2, \\ \text{Cov}(P_1, P_2, \sigma) &= (p - p^2)(k_1 - k_2)^2. \end{aligned} \quad (22)$$

Both of these covariance terms are positive for any possible value of p , so for both $\mathcal{F}(Q)$ and $\mathcal{F}(P)$ to be negative, it is necessary that $|\kappa| < 1$; if $|\kappa| > 1$, then either $1 + \kappa$ or $1 - \kappa$ will become negative, and in order for focusing to occur for both observables, this requires the corresponding covariance term must be negative, with the other being positive. We may position these two pure states a large distance away in phase space and allow for the covariances to keep increasing. Now, (15) and (16) become, respectively,

$$\begin{aligned} \mathcal{F}(Q) &= \frac{(1-\kappa)^2}{4} \mu^2 S \\ &+ \frac{(1-\kappa)}{\lambda} \mu \text{Cov}(Q_1, Q_2, \sigma_{12}) \left[\frac{1-\kappa}{4} \lambda \mu - 1 \right], \end{aligned} \quad (23)$$

$$\begin{aligned} \mathcal{F}(P) &= \frac{(1+\kappa)^2}{4} \lambda^2 R \\ &+ \frac{(1+\kappa)}{\mu} \lambda \text{Cov}(P_1, P_2, \sigma_{12}) \left[\frac{1+\kappa}{4} \lambda \mu - 1 \right]; \end{aligned} \quad (24)$$

here we see that joint focusing will occur if the covariance terms are made sufficiently large, and

$$\frac{\lambda \mu}{4} (1 + |\kappa|) < 1, \quad (25)$$

which is similar to (21).

Conclusion. In this Letter we extended the Arthurs-Kelly model to allow for probes prepared in an arbitrary state. In doing so, we showed that the resulting effective observable measured on our system is covariant under phase space translations. The marginals of these observables satisfy the error-disturbance relation, contrary to the claims of Di Lorenzo.

We showed Di Lorenzo's proposed measure of disturbance to actually be a measure of relative imprecision between two approximations of the ideal position or momentum observables. It is not a valid measure of disturbance, but does indicate the presence of focusing, where the marginals of a joint position and momentum measurement can be more precise than those performed sep-

arately. Focusing arises through the use of initial correlations between the probes, as was shown by examples with both entangled or separable probe states.

Acknowledgements. The authors thank Pieter Kok for numerous useful discussions up to and during the writing of this Letter. Further thanks are due to two anonymous referees, whose comments led to a number of clarifications in the presentation. The authors gratefully acknowledge support through the White Rose Studentship Network *Optimising Quantum Processes and Quantum Devices for future Digital Economy Applications*.

* tjb525@york.ac.uk

† paul.busch@york.ac.uk

- [1] E.H. Kennard. *Z. Phys.* **44**, 326 (1927).
- [2] H. Weyl. *Gruppentheorie und Quantenmechanik*. (Hirzel, Leipzig, 1928).
- [3] W. Heisenberg *Z. Phys.* **43**, 172 (1927).
- [4] M. Ozawa. *Ann. Phys.* **311**, 350 (2004).
- [5] C. Branciard. *PNAS* **110**, 6742 (2013).
- [6] A. Di Lorenzo. *Phys. Rev. Lett.* **110**, 120403 (2013).
- [7] P. Busch, P. Lahti, and R.F. Werner. *Phys. Rev. Lett.* **111**, 160405 (2013).
- [8] J. Dressel and F. Nori. *Phys. Rev. A* **89**, 022106 (2014).
- [9] Y. Watanabe, T. Sagawa, and M. Ueda. *Phys. Rev. A* **84**, 042121 (2011).
- [10] P. Busch. *Int. J. Theor. Phys.* **24**, 63 (1985).
- [11] P. Busch, T. Heinonen and P. Lahti. *Phys. Rep.* **452**, 155 (2007).
- [12] E. Arthurs and J.L. Kelly. *Bell Syst. Tech. J.* **44**, 725 (1965).
- [13] An early example of the use of arbitrary probe states in the context of the Arthurs-Kelly model is given by the paper [14]. To our knowledge, it was not until [6] that the possible implications of correlations between the probes were first considered.
- [14] D.M. Appleby. *J. Phys. A* **31**, 6419 (1998).
- [15] See supplemental material (which contains Refs. [1–4]) for the proof of Theorem 1, as well as the derivation of the marginal observables of $G^{(\lambda, \mu)}$, and the Ozawa [1] and BLW [2] errors for these observables.
- [16] C. Villani. *Optimal Transport: Old and New*. Springer, Berlin, 2009.
- [17] P. Busch, P. Lahti and R.F. Werner. *J. Math. Phys.* **55**, 042111 (2014).
- [18] A.S. Holevo. *Rep. Math. Phys.* **16**, 385 (1979).
- [19] R.F. Werner. *J. Math. Phys.* **25**, 1404 (1984).
- [20] G. Cassinelli, E. De Vito, and A. Toigo. *J. Math. Phys.* **44**, 4768 (2003).
- [21] J. Kiukas, P. Lahti, and K. Ylinen. *J. Math. Anal. Appl.* **319**, 783 (2006).
- [22] In the supplemental material for his paper, Di Lorenzo considers this particular example. The condition he gives for $\mathcal{F}(P)$ to be negative is in fact incorrect: using his notation, he gives the error-disturbance product $\epsilon_X^2 \eta_{K|X}^2 = (1 + 2r\tilde{\delta}_X/\delta_K)/(4 - 4r^2)$ (where $\eta_{K|X}^2 = \mathcal{F}(P)$), and so requires $\delta_K < 2\tilde{\delta}_X$ and $r \leq -\delta_K/2\tilde{\delta}_X$ to obtain the desired negativity. However, based on his given definitions, the product is equal to $(1 + 2r\delta_K/\tilde{\delta}_X)/(4 - 4r^2)$, and so the correct conditions are $\tilde{\delta}_X < 2\delta_K$ and $r \leq \tilde{\delta}_X/2\delta_K$.

SUPPLEMENTAL MATERIAL TO “FOCUSING IN ARTHURS-KELLY-TYPE JOINT MEASUREMENTS WITH CORRELATED PROBES”

Proof of Theorem 1

We first assume that the probes are prepared in the arbitrary pure state φ_{12} , with the position coordinate denoted by q_1 for the first probe and q_2 for the second. We further assume the considered system is prepared in the state ψ with position coordinate q . By using the form of the coupling unitary given in (4) and the identity $\exp(-i\lambda q P)\psi(x) = \psi(x - \lambda q)$, the combined state of the probes with the considered system is given by Ψ :

$$\begin{aligned}\Psi(q, q_1, q_2) &= U(\psi \otimes \varphi_{12})(q, q_1, q_2) \\ &= e^{-i\lambda Q P_1} e^{i\mu P Q_2} e^{-i\frac{\lambda\mu}{2}(\kappa+1)P_1 Q_2} (\psi \otimes \varphi_{12})(q, q_1, q_2) \\ &= e^{-i\lambda q P_1} e^{i\mu q_2 P} e^{-i\frac{\lambda\mu}{2}(\kappa+1)q_2 P_1} (\psi \otimes \varphi_{12})(q, q_1, q_2) \\ &= \psi(q + \mu q_2) \varphi_{12}(q_1 - \lambda q - \frac{\lambda\mu}{2}(\kappa+1)q_2, q_2).\end{aligned}\tag{26}$$

On the first probe we measure position, whilst on the second we measure momentum, so before we calculate the effective observable, we perform a Fourier transform on the final argument of Ψ :

$$\Psi(q, q_1, q_2) \mapsto \tilde{\Psi}(q, q_1, p_2) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dq_2 e^{-ip_2 q_2} \Psi(q, q_1, q_2).\tag{27}$$

The measurements performed on the probes are ideal measurements, described by projection-valued measures; the position measurement on the first probe is denoted by E^{Q_1} , and the momentum measurement on the second by E^{P_2} . As is given in (7), we find the effective observable on the considered system $G^{(\lambda, \mu)}$ via

$$\langle \psi | G^{(\lambda, \mu)}(X \times Y) \psi \rangle \equiv \langle \Psi | I \otimes E^{Q_1}(\lambda X) \otimes E^{P_2}(\mu Y) \Psi \rangle.\tag{28}$$

(The reason for the insertion of the scaling parameters λ, μ will become clear in (41) and (48).) By making use of (27), we calculate the form of $G^{(\lambda, \mu)}$:

$$\begin{aligned}\langle \psi | G^{(\lambda, \mu)}(X \times Y) \psi \rangle &= \frac{1}{2\pi} \int_{\mathbb{R}^8} dq dq' dq_1 dq'_1 dq_2 dq'_2 dp_2 dp'_2 e^{i(p'_2 q'_2 - p_2 q_2)} \overline{\psi(q' + \mu q'_2) \varphi_{12}(q'_1 - \lambda q' - \frac{\lambda\mu}{2}(\kappa+1)q'_2, q'_2)} \\ &\quad \times \psi(q + \mu q_2) \varphi_{12}(q_1 - \lambda q - \frac{\lambda\mu}{2}(\kappa+1)q_2, q_2) \langle q' | q \rangle \langle q'_1 | E^{Q_1}(\lambda X) q_1 \rangle \langle p'_2 | E^{P_2}(\mu Y) p_2 \rangle.\end{aligned}\tag{29}$$

After expressing $E^{Q_1}(\lambda X)$ and $E^{P_2}(\mu Y)$ in terms of pseudo-eigenvectors of Q_1 and P_2 , respectively

$$E^{Q_1}(\lambda X) = \int_{\lambda X} dq''_1 |q''_1\rangle \langle q''_1| = \lambda \int_X dq''_1 |\lambda q''_1\rangle \langle \lambda q''_1|;\tag{30a}$$

$$E^{P_2}(\mu Y) = \int_{\mu Y} dp''_2 |p''_2\rangle \langle p''_2| = \mu \int_Y dp''_2 |\mu p''_2\rangle \langle \mu p''_2|,\tag{30b}$$

the right hand side of (29) reduces to

$$\begin{aligned}\langle \psi | G^{(\lambda, \mu)}(X \times Y) \psi \rangle &= \frac{\lambda\mu}{2\pi} \int_{X \times Y} dq''_1 dp''_2 \int_{\mathbb{R}^4} dq dq' dq_2 dq'_2 e^{-i\mu p''_2(q_2 - q'_2)} \overline{\psi(q' + \mu q'_2) \varphi_{12}(\lambda(q''_1 - q' - \frac{\mu}{2}(\kappa+1)q'_2), q'_2)} \\ &\quad \times \psi(q + \mu q_2) \varphi_{12}(\lambda(q''_1 - q - \frac{\mu}{2}(\kappa+1)q_2), q_2) \langle q' | q \rangle \\ &= \int_{X \times Y} dq_1 dp_2 \left(\sqrt{\frac{\lambda\mu}{2\pi}} \int_{\mathbb{R}^2} dq dq_2 e^{-i\mu p_2 q_2} \psi(q + \mu q_2) \varphi_{12}(\lambda(q_1 - q - \frac{\mu}{2}(\kappa+1)q_2), q_2) |q\rangle \right)^* \\ &\quad \times \sqrt{\frac{\lambda\mu}{2\pi}} \int_{\mathbb{R}^2} dq dq_2 e^{-i\mu p_2 q_2} \psi(q + \mu q_2) \varphi_{12}(\lambda(q_1 - q - \frac{\mu}{2}(\kappa+1)q_2), q_2) |q\rangle.\end{aligned}\tag{31}$$

We define $q' = q + \mu q_2$, so $q_2 = \frac{1}{\mu}(q' - q)$, $dq'_2 = \frac{1}{\mu}dq'$ and $q + \frac{\mu}{2}(\kappa + 1)q_2 = \frac{1}{2}((1 - \kappa)q + (1 + \kappa)q')$. Therefore

$$\begin{aligned}
\langle \psi | \mathbf{G}^{(\lambda, \mu)}(X \times Y) \psi \rangle &= \int_{X \times Y} dq_1 dp_2 \left(\sqrt{\frac{\lambda}{2\pi\mu}} \int_{\mathbb{R}^2} dq dq' e^{ip_2(q-q')} \psi(q') \varphi_{12}(\lambda(q_1 - \frac{1}{2}((1 - \kappa)q + (1 + \kappa)q')), \frac{1}{\mu}(q' - q)) |q\rangle \right)^* \\
&\quad \times \sqrt{\frac{\lambda}{2\pi\mu}} \int_{\mathbb{R}^2} dq dq' e^{ip_2(q-q')} \psi(q') \varphi_{12}(\lambda(q_1 - \frac{1}{2}((1 - \kappa)q + (1 + \kappa)q')), \frac{1}{\mu}(q' - q)) |q\rangle \\
&= \int_{X \times Y} dq_1 dp_2 \left(\int_{\mathbb{R}^2} dq dq' K_{q_1 p_2}(q, q') \psi(q') |q\rangle \right)^* \int_{\mathbb{R}^2} dq dq' K_{q_1 p_2}(q, q') \psi(q') |q\rangle \\
&= \left\langle \psi \left| \left(\int_{X \times Y} dq_1 dp_2 K_{q_1 p_2}^* K_{q_1 p_2} \right) \psi \right. \right\rangle.
\end{aligned} \tag{32}$$

We have therefore found our effective observable:

$$\mathbf{G}^{(\lambda, \mu)}(X \times Y) = \int_{X \times Y} dq dp K_{qp}^* K_{qp}, \tag{33}$$

where K_{qp} has the kernel

$$K_{qp}(x, x') = e^{ip(x-x')} \varphi_{12}(\lambda(q - \frac{1}{2}((1 - \kappa)x + (1 + \kappa)x')), \frac{1}{\mu}(x' - x)). \tag{34}$$

K_{qp} can be rewritten in the following way:

$$\begin{aligned}
K_{qp} &= \int_{\mathbb{R}^2} dx dx' e^{ip(x-x')} \varphi_{12}(\lambda(q - \frac{1}{2}((1 - \kappa)x + (1 + \kappa)x')), \frac{1}{\mu}(x' - x)) |x\rangle \langle x'| \\
&= \int_{\mathbb{R}^2} dx dx' e^{ip((x-q)-(x'-q))} \varphi_{12}(\lambda(-\frac{1}{2}((1 - \kappa)(x - q) + (1 + \kappa)(x' - q))), \frac{1}{\mu}((x' - q) - (x - q))) |x\rangle \langle x'| \\
&= \int_{\mathbb{R}^2} dx dx' e^{ip(x-x')} \varphi_{12}(\lambda(-\frac{1}{2}((1 - \kappa)x + (1 + \kappa)x')), \frac{1}{\mu}(x' - x)) |x + q\rangle \langle x' + q| \\
&= e^{-iqP} \left(\int_{\mathbb{R}^2} dx dx' e^{ip(x-x')} \varphi_{12}(\lambda(-\frac{1}{2}((1 - \kappa)x + (1 + \kappa)x')), \frac{1}{\mu}(x' - x)) |x\rangle \langle x'| \right) e^{iqP} \\
&= e^{-iqP} e^{ipQ} \left(\int_{\mathbb{R}^2} dx dx' \varphi_{12}(\lambda(-\frac{1}{2}((1 - \kappa)x + (1 + \kappa)x')), \frac{1}{\mu}(x' - x)) |x\rangle \langle x'| \right) e^{-ipQ} e^{iqP} \\
&= W_{qp} K_{00} W_{qp}^*,
\end{aligned} \tag{35}$$

where $W_{qp} = \exp[-i(qP - pQ)]$ is the Weyl operator, which generates translations in phase space. From this we see

$$\begin{aligned}
W_{qp} \mathbf{G}^{(\lambda, \mu)}(Z) W_{qp}^* &= \int_Z dq' dp' W_{qp} K_{q' p'}^* K_{q' p'} W_{qp}^* \\
&= \int_Z dq' dp' (W_{qp} K_{q' p'} W_{qp}^*)^* (W_{qp} K_{q' p'} W_{qp}^*) \\
&= \int_{Z+(q,p)} dq' dp' K_{q' p'}^* K_{q' p'} \\
&= \mathbf{G}^{(\lambda, \mu)}(Z + (q, p)).
\end{aligned} \tag{36}$$

We have thus shown that if we start with probes prepared in an arbitrary pure state, the effective observable measured on the system is covariant under phase space translations.

Next, we consider the case of mixed states $\sigma = \sum_i p_i \sigma_i$, where the σ_i are arbitrary pure states. The post-coupling state is now given by $U(P_\psi \otimes \sigma)U^* = \sum_i p_i U(P_\psi \otimes \sigma_i)U^*$, and the effective observable is now found as follows:

$$\begin{aligned}
\langle \psi | \mathbf{G}^{(\lambda, \mu)}(X \times Y) \psi \rangle &= \text{tr} [U(P_\psi \otimes \sigma)U^* (I \otimes \mathbf{E}^{Q_1}(\lambda X) \otimes \mathbf{E}^{P_2}(\mu Y))] \\
&= \sum_i p_i \text{tr} [U(P_\psi \otimes \sigma_i)U^* (I \otimes \mathbf{E}^{Q_1}(\lambda X) \otimes \mathbf{E}^{P_2}(\mu Y))] \\
&= \left\langle \psi \left| \left(\sum_i p_i \mathbf{H}_i^{(\lambda, \mu)}(X \times Y) \right) \psi \right. \right\rangle,
\end{aligned} \tag{37}$$

where $H_i^{(\lambda, \mu)}$ is the covariant phase space observable associated with the probes prepared in the pure state σ_i . From this we quickly find that

$$\begin{aligned} W_{qp} G^{(\lambda, \mu)}(Z) W_{qp}^* &= \sum_i p_i W_{qp} H_i^{(\lambda, \mu)}(Z) W_{qp}^* \\ &= \sum_i p_i H_i^{(\lambda, \mu)}(Z + (q, p)) \\ &= G^{(\lambda, \mu)}(Z + (q, p)), \end{aligned} \quad (38)$$

proving the covariance, and hence Theorem 1.

The Marginals of $G^{(\lambda, \mu)}$

We find the marginals of the observable $G^{(\lambda, \mu)}$, $E^{(\lambda, \mu)}$ and $F^{(\lambda, \mu)}$, by integrating over the outcome space of the other variable (this may be seen as projecting down to a one-dimensional subspace of phase space):

$$E^{(\lambda, \mu)}(X) = G^{(\lambda, \mu)}(X \times \mathbb{R}); \quad (39a)$$

$$F^{(\lambda, \mu)}(Y) = G^{(\lambda, \mu)}(\mathbb{R} \times Y). \quad (39b)$$

Considering the case where the probes are prepared in the pure state φ_{12} , we first calculate $E^{(\lambda, \mu)}$:

$$\begin{aligned} E^{(\lambda, \mu)}(X) &= \int_{X \times \mathbb{R}} dq dp K_{qp}^* K_{qp} \\ &= \frac{\lambda}{\mu} \int_X dq \int_{\mathbb{R}^3} dx dx' dy' \left(\frac{1}{2\pi} \int_{\mathbb{R}} dp e^{ip(y' - x')} \right) \overline{\varphi_{12}(\lambda(q - \frac{1}{2}((1 - \kappa)x + (1 + \kappa)y')), \frac{1}{\mu}(y' - x))} \\ &\quad \times \varphi_{12}(\lambda(q - \frac{1}{2}((1 - \kappa)x + (1 + \kappa)x')), \frac{1}{\mu}(x' - x)) |y'\rangle\langle x'|. \\ &= \frac{\lambda}{\mu} \int_X dq \int_{\mathbb{R}^2} dx dx' \left| \varphi_{12}(\lambda(q - \frac{1}{2}((1 - \kappa)x + (1 + \kappa)x')), \frac{1}{\mu}(x' - x)) \right|^2 |x'\rangle\langle x'|, \end{aligned} \quad (40)$$

where we have used the identity $\int_{\mathbb{R}} dk \exp(ikx) = 2\pi\delta(x)$. We define $q' = x' - x$, so $x = x' - q'$, $dx = -dq'$ and $(1 - \kappa)x + (1 + \kappa)x' = 2x' - (1 - \kappa)q'$. $E^{(\lambda, \mu)}$ then takes the form:

$$\begin{aligned} E^{(\lambda, \mu)}(X) &= \frac{\lambda}{\mu} \int_X dq \int_{\mathbb{R}^2} dq' dx' \left| \varphi_{12}(\lambda(\frac{1}{2}(1 - \kappa)q' - (x' - q)), \frac{1}{\mu}q') \right|^2 |x'\rangle\langle x'| \\ &= \int_X dq \int_{\mathbb{R}} dx' e^{(\lambda, \mu)}(x' - q) |x'\rangle\langle x'| \\ &= \int_{\mathbb{R}} dq \chi_X(q) e^{(\lambda, \mu)}(Q - q) \\ &= (\chi_X * e^{(\lambda, \mu)})(Q), \end{aligned} \quad (41)$$

as is given in (9a). The probability distribution $e^{(\lambda, \mu)}$, which characterizes the noise in the measurement of $E^{(\lambda, \mu)}$, is of the form

$$e^{(\lambda, \mu)}(q) = \frac{\lambda}{\mu} \int_{\mathbb{R}} dq' \left| \varphi_{12}(\lambda(\frac{1}{2}(1 - \kappa)q' - q), \frac{1}{\mu}q') \right|^2, \quad (42)$$

with first and second moments

$$\begin{aligned}
e^{(\lambda, \mu)}[1] &= \int_{\mathbb{R}} dq q e^{(\lambda, \mu)}(q) \\
&= \frac{1}{\mu} \int_{\mathbb{R}^2} dq dq' \left(\frac{1}{2}(1 - \kappa)q' - \frac{1}{\lambda}q \right) \left| \varphi_{12}\left(q, \frac{1}{\mu}q'\right) \right|^2 \\
&= \int_{\mathbb{R}^2} dq dq' \left(\frac{\mu}{2}(1 - \kappa)q' - \frac{1}{\lambda}q \right) \left| \varphi_{12}(q, q') \right|^2 \\
&= \frac{\mu}{2}(1 - \kappa) \langle Q_2 \rangle_{\varphi_{12}} - \frac{1}{\lambda} \langle Q_1 \rangle_{\varphi_{12}};
\end{aligned} \tag{43}$$

$$\begin{aligned}
e^{(\lambda, \mu)}[2] &= \int_{\mathbb{R}} dq q^2 e^{(\lambda, \mu)}(q) \\
&= \int_{\mathbb{R}^2} dq dq' \left(\frac{\mu}{2}(1 - \kappa)q' - \frac{1}{\lambda}q \right)^2 \left| \varphi_{12}(q, q') \right|^2 \\
&= \frac{\mu^2}{4}(1 - \kappa)^2 \langle Q_2^2 \rangle_{\varphi_{12}} + \frac{1}{\lambda^2} \langle Q_1^2 \rangle_{\varphi_{12}} - \frac{\mu}{\lambda} \langle Q_1 Q_2 \rangle_{\varphi_{12}},
\end{aligned} \tag{44}$$

where $\langle Q_1 \rangle_{\varphi_{12}} = \langle \varphi_{12} | Q_1 \varphi_{12} \rangle$, etc. Using (43) and (44), the variance of $e^{(\lambda, \mu)}$ is

$$\text{Var}(e^{(\lambda, \mu)}) = \frac{1}{\lambda^2} \text{Var}(Q_1, \varphi_{12}) + \frac{\mu^2}{4} (1 - \kappa)^2 \text{Var}(Q_2, \varphi_{12}) - \frac{\mu}{\lambda} (1 - \kappa) \text{Cov}(Q_1, Q_2, \varphi_{12}), \tag{45}$$

where $\text{Cov}(Q_1, Q_2, \varphi_{12}) = \langle Q_1 Q_2 \rangle_{\varphi_{12}} - \langle Q_1 \rangle_{\varphi_{12}} \langle Q_2 \rangle_{\varphi_{12}}$ is the covariance of Q_1 and Q_2 with respect to φ_{12} . In a similar fashion, we derive an explicit form for $F^{(\lambda, \mu)}$, the first step of which is to perform a Fourier transform on φ_{12} :

$$\varphi_{12}\left(\lambda\left(q - \frac{1}{2}((1 - \kappa)x + (1 + \kappa)x')\right), \frac{1}{\mu}(x' - x)\right) = \frac{1}{2\pi} \int_{\mathbb{R}^2} dw dz e^{iw\left(q - \frac{1}{2}((1 - \kappa)x + (1 + \kappa)x')\right)} e^{iz(x' - x)} \tilde{\varphi}_{12}\left(\lambda w, \frac{z}{\mu}\right), \tag{46a}$$

$$\overline{\varphi_{12}\left(\lambda\left(q - \frac{1}{2}((1 - \kappa)x + (1 + \kappa)y')\right), \frac{1}{\mu}(y' - x)\right)} = \frac{1}{2\pi} \int_{\mathbb{R}^2} dw' dz' e^{-iw'\left(q - \frac{1}{2}((1 - \kappa)x + (1 + \kappa)y')\right)} e^{-iz'(y' - x)} \overline{\tilde{\varphi}_{12}\left(\lambda w', \frac{z'}{\mu}\right)}, \tag{46b}$$

and so we find

$$\begin{aligned}
F^{(\lambda, \mu)}(Y) &= \frac{\lambda}{2\pi\mu} \int_{\mathbb{R} \times Y} dq dp \int_{\mathbb{R}^3} dx dx' dy' e^{ip(y' - x')} \int_{\mathbb{R}^4} \frac{dw dw' dz dz'}{4\pi^2} e^{iq(w - w')} e^{-\frac{iw}{2}((1 - \kappa)x + (1 + \kappa)x')} \\
&\quad \times e^{\frac{iw'}{2}((1 - \kappa)x + (1 + \kappa)y')} e^{ix(z' - z)} e^{izx'} e^{-iz'y'} \tilde{\varphi}_{12}\left(\lambda w, \frac{z}{\mu}\right) \overline{\tilde{\varphi}_{12}\left(\lambda w', \frac{z'}{\mu}\right)} |y'\rangle \langle x'| \\
&= \frac{\lambda}{2\pi\mu} \int_Y dp \int_{\mathbb{R}^3} dx dx' dy' e^{ip(y' - x')} \int_{\mathbb{R}^4} \frac{dw dw' dz dz'}{2\pi} \left(\frac{1}{2\pi} \int_{\mathbb{R}} dq e^{iq(w - w')} \right) e^{-\frac{iw}{2}((1 - \kappa)x + (1 + \kappa)x')} \\
&\quad \times e^{\frac{iw'}{2}((1 - \kappa)x + (1 + \kappa)y')} e^{ix(z' - z)} e^{izx'} e^{-iz'y'} \tilde{\varphi}_{12}\left(\lambda w, \frac{z}{\mu}\right) \overline{\tilde{\varphi}_{12}\left(\lambda w', \frac{z'}{\mu}\right)} |y'\rangle \langle x'| \\
&= \frac{\lambda}{2\pi\mu} \int_Y dp \int_{\mathbb{R}^3} dw dz dz' \int_{\mathbb{R}^2} dx' dy' e^{i\left(p + \frac{w}{2}(1 + \kappa)\right)(y' - x')} \left(\frac{1}{2\pi} \int_{\mathbb{R}} dx e^{ix(z' - z)} \right) \\
&\quad \times e^{izx'} e^{-iz'y'} \tilde{\varphi}_{12}\left(\lambda w, \frac{z}{\mu}\right) \overline{\tilde{\varphi}_{12}\left(\lambda w, \frac{z'}{\mu}\right)} |y'\rangle \langle x'| \\
&= \frac{\lambda}{2\pi\mu} \int_Y dp \int_{\mathbb{R}^2} dx' dy' \int_{\mathbb{R}^2} dw dz e^{i\left(p + \frac{w}{2}(1 + \kappa) - z\right)(y' - x')} \left| \tilde{\varphi}_{12}\left(\lambda w, \frac{z}{\mu}\right) \right|^2 |y'\rangle \langle x'| \\
&= \frac{\lambda}{\mu} \int_Y dp \int_{\mathbb{R}^2} dw dz \left| \tilde{\varphi}_{12}\left(\lambda w, \frac{z}{\mu}\right) \right|^2 \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dy' e^{iy'\left(p + \frac{w}{2}(\kappa + 1) - z\right)} |y'\rangle \right) \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dx' e^{ix'\left(p + \frac{w}{2}(\kappa + 1) - z\right)} |x'\rangle \right)^* \\
&= \frac{\lambda}{\mu} \int_Y dp \int_{\mathbb{R}^2} dw dz \left| \tilde{\varphi}_{12}\left(\lambda w, \frac{z}{\mu}\right) \right|^2 \left| p + \frac{w}{2}(\kappa + 1) - z \right\rangle \left\langle p + \frac{w}{2}(\kappa + 1) - z \right|.
\end{aligned} \tag{47}$$

Defining $p' = p + \frac{w}{2}(\kappa + 1) - z$, so $z = p - p' + \frac{w}{2}(\kappa + 1)$ and $dz = -dp'$, $F^{(\lambda, \mu)}$ takes the form

$$\begin{aligned} F^{(\lambda, \mu)}(Y) &= \frac{\lambda}{\mu} \int_Y dp \int_{\mathbb{R}^2} dw dp' \left| \tilde{\varphi}_{12}(\lambda w, \frac{1}{\mu}(p - p' + \frac{w}{2}(\kappa + 1))) \right|^2 |p'\rangle \langle p'| \\ &= \int_Y dp \int_{\mathbb{R}} dp' f^{(\lambda, \mu)}(p' - p) |p'\rangle \langle p'| \\ &= \int_{\mathbb{R}} dp \chi_Y(p) f^{(\lambda, \mu)}(P - p) \\ &= (\chi_Y * f^{(\lambda, \mu)})(P), \end{aligned} \quad (48)$$

as is given in (9b). The probability distribution $f^{(\lambda, \mu)}$ is of the form

$$f^{(\lambda, \mu)}(p) = \frac{\lambda}{\mu} \int_{\mathbb{R}} dw \left| \tilde{\varphi}_{12}(\lambda w, \frac{1}{\mu}(\frac{w}{2}(\kappa + 1) - p)) \right|^2. \quad (49)$$

By using the identity

$$\tilde{\varphi}_{12}(\lambda p, \frac{1}{\mu}p') = \frac{\mu}{\lambda} \tilde{\varphi}_{12}(\frac{1}{\lambda}p, \mu p'), \quad (50)$$

and following the same method used to derive (43) and (44), we find the first and second moments of $f^{(\lambda, \mu)}$:

$$f^{(\lambda, \mu)}[1] = \frac{\lambda}{2}(1 + \kappa) \langle P_1 \rangle_{\varphi_{12}} - \frac{1}{\mu} \langle P_2 \rangle_{\varphi_{12}}; \quad (51)$$

$$f^{(\lambda, \mu)}[2] = \frac{\lambda^2}{4}(1 + \kappa)^2 \langle P_1^2 \rangle_{\varphi_{12}} + \frac{1}{\mu^2} \langle P_2^2 \rangle_{\varphi_{12}} - \frac{\lambda}{\mu}(1 + \kappa) \langle P_1 P_2 \rangle_{\varphi_{12}}. \quad (52)$$

From these, the variance of $f^{(\lambda, \mu)}$ is given by

$$\text{Var}(f^{(\lambda, \mu)}) = \frac{\lambda^2}{4}(1 + \kappa)^2 \text{Var}(P_1, \varphi_{12}) + \frac{1}{\mu^2} \text{Var}(P_2, \varphi_{12}) - \frac{\lambda}{\mu}(1 + \kappa) \text{Cov}(P_1, P_2, \varphi_{12}). \quad (53)$$

Remark 1. It is now clear why we use the scaled sets λX and μY in (7) and (28): the scaling is such that the marginal observables $E^{(\lambda, \mu)}$ and $F^{(\lambda, \mu)}$ are direct smearings of position and momentum, rather than of scaled versions.

We will now return to the case of the mixed state $\sigma = \sum_i p_i \sigma_i$, where $\sigma_i = P_{\varphi_i}$ are arbitrary pure states. The marginals of the effective observable $G^{(\lambda, \mu)}$ are now given in terms of the marginals of the effective observables derived from σ_i :

$$\begin{aligned} E^{(\lambda, \mu)}(X) &= G^{(\lambda, \mu)}(X \times \mathbb{R}) = \sum_i p_i H_i^{(\lambda, \mu)}(X \times \mathbb{R}) = \sum_i p_i M_i^{(\lambda, \mu)}(X) \\ &= \sum_i p_i (\chi_X * m_i^{(\lambda, \mu)})(Q) = (\chi_X * e^{(\lambda, \mu)})(Q), \end{aligned} \quad (54a)$$

$$\begin{aligned} F^{(\lambda, \mu)}(Y) &= G^{(\lambda, \mu)}(\mathbb{R} \times Y) = \sum_i p_i H_i^{(\lambda, \mu)}(\mathbb{R} \times Y) = \sum_i p_i N_i^{(\lambda, \mu)}(Y) \\ &= \sum_i p_i (\chi_Y * n_i^{(\lambda, \mu)})(P) = (\chi_Y * f^{(\lambda, \mu)})(P). \end{aligned} \quad (54b)$$

As such, these marginals again have the form (9a), (9b), with the probability distributions

$$e^{(\lambda, \mu)}(q) = \sum_i p_i m_i^{(\lambda, \mu)}(q) = \frac{\lambda}{\mu} \sum_i p_i \int_{\mathbb{R}} dq' \left| \varphi_i(\lambda(\frac{1}{2}(1 - \kappa)q' - q), \frac{1}{\mu}q') \right|^2, \quad (55a)$$

$$f^{(\lambda, \mu)}(p) = \sum_i p_i n_i^{(\lambda, \mu)}(p) = \frac{\lambda}{\mu} \sum_i p_i \int_{\mathbb{R}} dw \left| \tilde{\varphi}_i(\lambda w, \frac{1}{\mu}(\frac{w}{2}(\kappa + 1) - p)) \right|^2. \quad (55b)$$

From here, and by using equations (43), (44), (51), (52), the first and second moments of these distributions can be readily calculated:

$$\begin{aligned} e^{(\lambda, \mu)}[1] &= \int_{\mathbb{R}} dq q e^{(\lambda, \mu)} = \sum_i p_i \int_{\mathbb{R}} dq q m_i^{(\lambda, \mu)}(q) = \sum_i p_i m^{(\lambda, \mu)}[1] \\ &= \sum_i p_i \left(\frac{\mu}{2} (1 - \kappa) \langle Q_2 \rangle_{\varphi_i} - \frac{1}{\lambda} \langle Q_1 \rangle_{\varphi_i} \right) \end{aligned} \quad (56)$$

$$\begin{aligned} &= \frac{\mu}{2} (1 - \kappa) \langle Q_2 \rangle_{\sigma} - \frac{1}{\lambda} \langle Q_1 \rangle_{\sigma}, \\ e^{(\lambda, \mu)}[2] &= \frac{\mu^2}{4} (1 - \kappa)^2 \langle Q_2^2 \rangle_{\sigma} + \frac{1}{\lambda^2} \langle Q_1^2 \rangle_{\sigma} - \frac{\mu}{\lambda} \langle Q_1 Q_2 \rangle_{\sigma}, \end{aligned} \quad (57)$$

$$f^{(\lambda, \mu)}[1] = \frac{\lambda}{2} (1 + \kappa) \langle P_1 \rangle_{\sigma} - \frac{1}{\mu} \langle P_2 \rangle_{\sigma}, \quad (58)$$

$$f^{(\lambda, \mu)}[2] = \frac{\lambda^2}{4} (1 + \kappa)^2 \langle P_1^2 \rangle_{\sigma} + \frac{1}{\mu^2} \langle P_2^2 \rangle_{\sigma} - \frac{\lambda}{\mu} (1 + \kappa) \langle P_1 P_2 \rangle_{\sigma}, \quad (59)$$

and so these probability distributions have variances of the form given in (45) and (53)

$$\text{Var}(e^{(\lambda, \mu)}) = \frac{1}{\lambda^2} \text{Var}(Q_1, \sigma) + \frac{\mu^2}{4} (1 - \kappa)^2 \text{Var}(Q_2, \sigma) - \frac{\mu}{\lambda} (1 - \kappa) \text{Cov}(Q_1, Q_2, \sigma), \quad (60)$$

$$\text{Var}(f^{(\lambda, \mu)}) = \frac{\lambda^2}{4} (1 + \kappa)^2 \text{Var}(P_1, \sigma) + \frac{1}{\mu^2} \text{Var}(P_2, \sigma) - \frac{\lambda}{\mu} (1 + \kappa) \text{Cov}(P_1, P_2, \sigma). \quad (61)$$

Note that even if we had specified that the pure states σ_i were product states, we would still find the covariance terms appearing as a result of the classical correlations between them.

Error values for $E^{(\lambda, \mu)}$ and $F^{(\lambda, \mu)}$

In this section, we briefly review the definitions of error presented by Ozawa [1] and Busch, Lahti and Werner [2]. While these definitions take different approaches, the purpose of this section is to show explicitly that in the case of observables of the form Q_m , where $Q_m(X) = (\chi_X * m)(Q)$ with m a probability distribution, these error measures actually coincide. This result applies, in particular, to $E^{(\lambda, \mu)}$ and $F^{(\lambda, \mu)}$. (It seems useful to note this observation for future reference; the result is implicit from calculations of these quantities found in various places in the literature.)

We shall begin by discussing the definition of error given by Ozawa [1]. In its most general terms, Ozawa's definition is an attempt to generalise root-mean-square deviations for quantum observables. Consider a system in a state ψ , upon which one wishes to measure the observable E^A with first moment operator $A = \int_{\mathbb{R}} x dE^A(x)$. If we couple this to an auxiliary system, described by the Hilbert space \mathcal{K} in a state ξ , with a coupling unitary U , and perform a sharp pointer measurement Z with first moment Z , then the error is given by

$$\epsilon(E, E^A, \psi)^2 = \langle \psi, \xi | (U^*(I \otimes Z)U - A \otimes I)^2 | \psi, \xi \rangle, \quad (62)$$

where E is the effective observable given by

$$E(X) = \text{tr}_{\mathcal{K}} [U^*(I \otimes Z(X))U(I \otimes P_{\xi})]. \quad (63)$$

This can be expressed in terms of quantities pertaining to the measured system alone:

$$\epsilon(E, E^A, \psi)^2 = \langle \psi | (E[1] - A)^2 | \psi \rangle + \langle \psi | (E[2] - E[1]^2) | \psi \rangle. \quad (64)$$

With (64) at hand, we now return to the case we wish to consider, namely the error of Q_m with respect to the ideal measurement E^Q , $\epsilon(Q_m, E^Q, \psi)$. It is easily shown that the first and second moment operators of the observable Q_{μ} are given by

$$Q_m[1] = Q + m[1], \quad Q_m[2] = Q_m[1]^2 + \text{Var}(m). \quad (65)$$

This gives

$$\epsilon(Q_m, E^Q, \psi)^2 = m[2] = m[1]^2 + \text{Var}(m). \quad (66)$$

We next recall the definition of error introduced in [2] as an operationally meaningful quantum version of root-mean-square error. For any two probability measures α, β on \mathbb{R} a *coupling* is defined to be a probability measure γ on $\mathbb{R} \times \mathbb{R}$ with α and β as the Cartesian marginals. The set of couplings between α and β will be denoted $\Gamma(\alpha, \beta)$. Then, the (Wasserstein) 2-distance [3] of α and β is defined as

$$\mathcal{D}_2(\alpha, \beta) = \inf_{\gamma \in \Gamma(\alpha, \beta)} \mathcal{D}_2^\gamma(\alpha, \beta) = \inf_{\gamma \in \Gamma(\alpha, \beta)} \left(\int |x - y|^2 d\gamma(x, y) \right)^{\frac{1}{2}} \quad (67)$$

The existence of an optimal coupling is known, see [3, Theorem 4.1], but it does not imply that $\mathcal{D}_2(\alpha, \beta)$ is finite. This is a distance between probability measures due to the choice of the minimizing joint probability.

We can now define the (*Wasserstein*) 2-distance between observables E, F on \mathbb{R} using the notation p_ρ^E, p_ρ^F for their probability measures with respect to the state ρ :

$$\Delta_2(E, F) := \sup_{\rho} \mathcal{D}_2(p_\rho^E, p_\rho^F).$$

As a direct application of [4, Lemma 7], one obtains

$$\Delta_2(Q_m, E^Q)^2 = \mathcal{D}_2(m, \delta_0)^2 = m[2] \equiv \epsilon(Q_m, E^Q, \psi)^2. \quad (68)$$

It is a remarkable that $\epsilon(Q_m, E^Q, \psi)$ coincides with $\Delta_2(Q_m, E^Q)$ considering that the former is expressly defined as a state-specific quantity while the latter represents a worst-case error measure across all states. This coincidence underpins the intuitive idea that the smearing measure m characterizes the random and systematic errors in a measurement of the observable Q_m as an approximation of E^Q .

* tjb525@york.ac.uk

† paul.busch@york.ac.uk

[1] M. Ozawa. *Ann. Phys.* **311**, 350 (2004).

[2] P. Busch, P. Lahti, and R.F. Werner. *Phys. Rev. Lett.* **111**, 160405 (2013).

[3] C. Villani. *Optimal Transport: Old and New*. Springer, Berlin, 2009.

[4] P. Busch, P. Lahti and R.F. Werner. *J. Math. Phys.* **55**, 042111 (2014).